DOCUMENT RESUME

ED 260 046 SP 026 347

AUTHOR Vickers, R. R., Jr.
TITLE Side Effects of Physical Training: Association of Fitness Improvement to Espirt de Corps, Performance, Health, and Attrition in Marine Corps Basic Training. Report No. 83-37.
INSTITUTION Naval Health Research Center, San Diego, Calif.
PUB DATE Dec 83
NOTE 32p.; For related document, see SP 026 348.
PUB TYPE Reports - Research/Technical (143)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Military Personnel; Military Training; Performance Factors; *Physical Fitness; Physical Health; *Program Effectiveness; *Program Evaluation; *Training Methods
IDENTIFIERS *Marine Corps

ABSTRACT

Physical training is a significant element of Marine Corps basic training which may affect nonfitness basic training outcomes in addition to improving fitness. If so, physical training side effects should be considered when designing and evaluating physical training programs. This study capitalized on naturally occurring platoon differences in fitness outcomes to classify basic training platoons into relatively high and low improvement groups. The high improvement group had better attitudes toward the Marine Corps, greater perceived self-improvement, and performed better during basic training. The groups had comparable illness rates during basic training, but there was more medical attrition and overall attrition in the high improvement group. The groups did not differ with regard to Fleet Marine Force promotion or attrition rates. These findings demonstrate that physical training programs may indeed influence nonfitness basic training outcomes. Additional research is needed to verify that physical training differences actually cause nonfitness basic training outcomes, to determine whether these side effects are large enough to be of practical importance, and to identify specific aspects of physical training programs that are important for promoting nonfitness basic training objectives. (Author)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED260046

SIDE EFFECTS [illegible]

[illegible]

R. R. VICKERS, JR.

REPORT NO. [illegible]

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

This document has been reproduced as received from the person or organization originating it.

Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

NAVAL [illegible] LABORATORY COMMAND
[illegible]

SIDE EFFECTS OF PHYSICAL TRAINING:

Association of Fitness Improvement to Esprit de Corps, Performance, Health, and Attrition in Marine Corps Basic Training[†]

Ross R. Vickers, Jr.

Health Psychology Department

Naval Health Research Center

P. O. Box 85122

San Diego, CA. 92138-9174

[†]Report Number 83-37, supported by the Naval Medical Research and Development Command, Department of the Navy, under research Work Unit M0096-PN.001-1035. The views presented are those of the authors. No endorsement by the Department of the Navy has been given or should be inferred.

SUMMARY

Physical training (PT) is an important component of Marine Corps basic training (BT). The primary PT objective is to improve recruit fitness to prepare the men for the rigors of the Fleet Marine Force (FMF). However, because PT is part of an integrated program designed to achieve a variety of outcomes, it may be important to consider possible side effects of PT on nonfitness outcomes when designing or evaluating PT programs.

Side effects of PT were investigated by relating naturally occurring platoon differences in fitness outcomes to (a) Esprit de Corps (measured by attitudes toward the Marine Corps and perceptions of BT leadership and group cohesion), (b) improved self-confidence (measured by perceived self-improvement during BT), (c) acquisition of military skills and knowledge (measured by BT tests and Drill Instructor ratings), (d) BT health (measured by illness incidents requiring 1 or more dispensary visits), (e) BT attrition, and (f) FMF success (measured by attrition and promotion rates following BT). The data for the analyses were originally collected as part of three BT stress studies carried out between 1978 and 1981.

The scores from the standard physical fitness test administered at the end of the first two weeks of BT and again in the tenth week were employed to determine fitness improvement for individual recruits. Platoons were classified as low or high improvement based on the average improvement of the platoon members. Both simple difference scores and residualized gain scores were used to estimate improvement to minimize the risk of erroneous classifications resulting from statistical limitations of either score used alone. The relationships between PT improvement and the outcomes described above were determined by t-tests between the two groups and contingency table analysis.

The high improvement group showed (a) more positive attitudes toward the Marine Corps and Marine Corps philosophy (i.e., greater acceptance of authority, higher commitment to doing well in the Marine Corps, and greater general satisfaction with the Marine Corps); (b) greater feelings of self-improvement during BT; and (c) better performance during BT. Another positive correlate of high improvement was that recruits in this group described their Drill Instructors as better examples of what a marine should be. A possible negative effect of PT was higher attrition in the high improvement group; the most consistent element of the attrition pattern was higher medical attrition in the high improvement group.

The study objective was to determine whether PT might have side effects which should be taken into account when designing PT programs. The significant associations described above show that PT side effects may be important. Additional research is needed to determine whether PT actually causes the nonfitness outcomes or is correlated with those outcomes for some other reason, whether the side effects are large enough to be of practical concern, and what specific aspects of PT programs affect which outcomes. This information would provide a basis for guidelines to design PT programs to enhance the overall effects of PT.

TABLE OF CONTENTS

Summary.....i

Table of Contents.....iii

Acknowledgments.....iv

Introduction.....1

Method.....3
- Samples.....3
- Physical Training Improvement Classifications.....4
- Esprit de Corps Measures.....4
- Self-Improvement Scales.....7
- Performance Measures.....7
- Health Indicators.....8
- FMF Success.....8
- Analysis Procedures.....9

Results.....10
- Hypothesis 1: High fitness improvement will be associated with greater Esprit de Corps.....10
- Hypothesis 2: HIgh fitness improvement will be associated with greater perceived self-improvement in recruits.....11
- Hypothesis 3: Fitness improvement level will be related to BT performance.....11
- Hypothesis 4: High fitness improvement will be associated with poorer health in BT.....11
- Hypothesis 5: Attrition during basic training will be higher in the high improvement group.....12
- Hypothesis 6: Graduates from the high improvement group will have greater FMF success than those from the low improvement group.....13

Discussion and Conclusions.....13

References.....16

Appendices

Appendix A: Defining Fitness Improvement Groups.....A-1
- Sample Comparisons.....A-2
- Platoon Differences in Fitness Improvement.....A-3
- Tests of Regression Assumptions for Gain Scores.....A-4
- Effectiveness of Group Classification.....A-6

ACKNOWLEDGMENTS

This report has benefited substantially from comments and suggestions on earlier drafts offered by Dr. James A. Hodgdon. The assistance of Ms. Linda K. Hervig in the performance of some of the analyses also contributed substantially to the preparation of the report.

INTRODUCTION

Improving the physical fitness of recruits to prepare them for the physical rigors of Fleet Marine Force (hereafter, FMF) activities is one objective of Marine Corps basic training (hereafter, BT) (1, p.1-1). The success of the current physical training program (hereafter, PT) is evident; average scores on a standard physical fitness test increase by 25% during BT (2). Even though this improvement indicates success in achieving the primary PT goal, there may be times when program modifications are considered. For example, the program might be restructured if changes would make more time available for higher priority training activities or would make BT graduates better prepared for the specific physical requirements of the FMF. The present report examines six possible side effects of PT which may deserve attention when considering possible program modifications.

The side effect concept for PT is borrowed from medical terminology. In medical terms, a side effect is a physical or psychological reaction to treatment which is not the primary or intended treatment goal. Because recruit PT is part of an integrated program designed to achieve multiple BT objectives, a PT program may produce side effects by influencing nonfitness BT outcomes, including recruits' acquisition of self-discipline, military skills, marksmanship, military bearing, and Esprit de Corps (1, p. 1-1).

Incidental data analyses in our earlier work on the effects of psychological stress in BT (2) showed substantial platoon differences in fitness at the end of BT. These differences occurred despite the fact that the PT program was highly structured. However, training procedures permit some augmentation of the standard program at the discretion of the training personnel (up to carefully specified limits). The observed platoon differences implied that differential use of this discretionary latitude resulted in significantly different PT programs in different platoons.

The naturally occurring platoon fitness differences provided the opportunity to explore possible side effects of PT programs. This report therefore examines how platoons which differed in fitness improvement during BT fared with respect to other BT outcomes. Platoons with relatively large fitness improvements are compared to platoons with relatively small improvements to test six hypotheses:

Hypothesis 1: High fitness improvement will be associated with greater Esprit de Corps.

Difficult initiations produce positive attitudes toward the organization joined (3-5), so high improvement should be associated with more positive attitudes toward the Marine Corps at the end of BT. This basic hypothesis was extended to include increased group cohesion and more positive perceptions of leadership because positive evaluation may generalize to perceptions of other members of the organization.

Hypothesis 2: High fitness improvement will be associated with greater perceived self-improvement in recruits.

Interviews with recruits suggested that feelings of self-improvement arise from mastering significant BT challenges. PT is one of the challenges that recruits must master, so more difficult PT should produce greater gains in self-esteem. If improved physical capabilities contribute to improved self-esteem, high improvement PT programs can promote the BT objective of developing recruits' self-reliance and confidence (1, p. 2-2).

Hypothesis 3: Fitness improvement will be related to BT performance.

This hypothesis concerns the BT objectives of developing military skills, marksmanship, and military bearing. These performance variables are focal points for the training program (1, p. 1-1), but two conflicting possibilities appeared reasonable. Programs that produce larger than average improvement could produce better performance as part of a general striving for excellence. However, if fitness training is emphasized at the expense of other aspects of training, this emphasis would impair performance in other areas.

The three remaining hypotheses concern possible side effects which were not explicit BT objectives, but which are logically related to those objectives. In contrast to the first three hypotheses which all reflect potential positive side effects, two of the remaining hypotheses concern possible negative side effects.

Hypothesis 4: High fitness improvement will be associated with poorer health during BT.

More demanding physical programs probably will produce more sprains, strains, and other minor injuries because the increased demands push more recruits to their physical limits and beyond. A more speculative possibility is that heavy physical demands will reduce resistance to infectious diseases which are common in BT (6,7). These possible effects are important because illness is inherently undesirable and likely to impede progress toward stated BT objectives.

Hypothesis 5: High improvement will be associated with higher attrition.

This hypothesis assumes that high improvement programs will cause more injuries, exacerbate more old medical problems, and/or cause more recruits to give up because they feel they cannot meet the strict program requirements. The first two possibilities would lead to higher rates of medical attrition and the third possibility to increased attrition for behavioral causes.

Hypothesis 6: Graduates from the high improvement group will have greater FMF success than those from the low improvement group.

All other things being equal, FMF success implies that a recruit satisfied the BT objective of learning to maintain or improve those traits that distinguish a Marine (1, p. 1-1). However, the usefulness of FMF success as a general criterion of Marine Corps success is limited by the qualification that FMF conditions are probably important contributors to this outcome. Thus, this hypothesized association would be expected to be weak.

METHOD

Samples

Data from three studies originally designed to assess the effects of psychological stresses on BT outcomes were employed to test the hypotheses outlined above. Descriptive statistics for the recruit samples in each study are provided in Table 1.

TABLE 1

DESCRIPTION OF RECRUIT COHORTS AND STUDY DESIGNS

STUDY	SAMPLE	AGE	EDUCATION	GCT	RACE[a]			
					W	B	H	O
1	413	18.7	11.4	104.1	73.9%	11.1%	11.4%	3.7%
2(a)[b]	2360	19.4	11.5	103.2	***	***	***	***
2(b)	425	19.1	11.6	104.8	65.9%	15.3%	14.1%	4.6%
3	2648	18.9	11.7	104.1	70.7%	16.3%	7.7%	5.5%

[a]Race groups were W = White, B = Black, H = Hispanic, and O = Other. Percentages were based on responses to a questionnaire item concerning race. Nonrespondents comprised 12.8% of the Study 1 sample, 13.7% of Study 2, and 2.4% of Study 3. The first two studies explicitly noted that providing this information was optional; the third study made no comment regarding this specific item.

[b]In Study 2 only the subset of 425 recruits completed the questionnaires. Other data were obtained for the full sample.

***Respondents in Study 2(a) did not complete the background questionnaire that provided race information.

Physical Training Improvement Classifications

Physical fitness scores were the results of a standard BT physical fitness test consisting of pull-ups, sit-ups, and a 3-mile run. Scores on each test component could range from 0 to 100 for a total possible score of 300. These tests were administered during the third week of training and again seven weeks later in the last week of training (hereafter, PFT1 and PFT2).

Two physical fitness improvement scores were computed for each recruit. First, the difference between the scores on the first and second test was computed for each recruit. Second, regression analysis produced an equation to predict PFT2 from PFT1. Each recruit's predicted value was then subtracted from his actual final score to measure residualized gain (hereafter, gain).

Platoons were classified as high improvement if they were in the top 40% of the platoons in their study based on (a) average difference, (b) average gain, (c) median difference, and (d) median gain. Low improvement platoons were those in the lowest 40% of the sample for all four criteria. The analyses were restricted to extreme groups to ensure clear-cut group differences in fitness improvement.

Multiple criteria ensured that classification was based on unbiased estimates of improvement that were independent of initial fitness level and that were not influenced by one or two extreme scores in a platoon. Difference scores are unbiased estimates of improvement, but can be influenced by initial fitness level (8). Gain scores are independent of initial fitness level, but are subject to bias when based on inappropriate assumptions about growth curves (8,9). The available data did not provide sufficient information to define growth curves precisely, so difference scores were used to ensure that the potential bias did not affect the findings. Finally, average scores are sensitive to outliers, but median scores are not (10, pp. 61-62). The combined criteria therefore protected against important sources of classification error when defining the fitness improvement groups. Details of a series of analyses supporting the classification process are given in Appendix A.

Esprit de Corps Measures

Esprit de Corps was assessed by measuring attitudes toward the Marine Corps, perceptions of leadership in BT, and perceptions of group cohesion in BT. Each of the measures represents an evaluation of the Marine Corps or perception of fellow Marines consistent with the general definition of Esprit given in BT manuals (1, p. 1-1). Each attitude or perception was measured by a multi-item Likert scale taken from questionnaires employed in Studies 1 and 2 (2,11,12).

The four attitudes measured were:

Affiliation: The extent to which the recruit identifies himself more with the Marine Corps than with the civilian population.

Example Items: (a) I feel that my outlook is really more that of a civilian than a Marine. (Reverse scored)

(b) If my Commanding Officer offered me an honorable discharge right now, I would not take it.

Source: Booth and Hoiberg (13).

Authority: The extent to which the recruit possesses traditional views of authority including acceptance of the necessity for and importance of accepting authority in the Marine Corps.

Example Items: (a) The discipline you get in the Marine Corps is good for you.

(b) A Marine should not be allowed to talk back to his superiors.

Source: Booth and Hoiberg (13).

Commitment: The extent to which the recruit feels that it is important to him personally to achieve and maintain a high level of performance in the Marine Corps.

Example Items: (a) It is important to me personally to have a good record in the Marine Corps.

(b) I don't care how well I do in the Marine Corps. (Reverse scored)

Source: Drucker (14).

Satisfaction: The extent to which the recruit holds a positive overall evaluation of the Marine Corps and perceives it as better than alternative occupations.

Example Items: (a) All in all, I am very satisfied with being a Marine.

(b) If I had my choice between joining the Navy, Army, Air Force, or Marines, I would still prefer to join the Marines.

Source: Modified from Quinn and Shepard (15).

The leadership measures were four scales that were present in both Study 1 and Study 2. Additional leadership scales available in one study but not the other were excluded from the analyses to ensure that conclusions were based on reliable differences. The four scales present in both studies included the best marker variables for the two leadership factors in our questionnaire (12), so the major BT leadership elements perceived by recruits were reflected despite the restriction to four scales.

Leader Structure: The extent to which Drill Instructors provided means-end structuring by detailing who was to do what and when.

Example Items: (a) Our Drill Instructors told us exactly how to do things.

(b) Drill Instructors told us why things had to be done.

Leader Support: The extent to which Drill Instructors communicated a concern for the well-being of the recruits and a respect for the platoon.

Example Items: (a) The Drill Instructors were interested in our welfare.

(b) The Drill Instructors were proud of the platoon and the recruits in it.

Referent Power: The extent to which Drill Instructors were regarded as setting a good example which the recruits want to copy.

Example Items: (a) I would like to be like my Drill Instructors.

(b) I respect my Drill Instructors as people.

Expert Power: The extent to which Drill Instructors were expert and knowledgeable in their job.

Example Items: (a) My Drill Instructors are well-qualified for their jobs.

(b) My Drill Instructors are very good at what they do.

Group cohesion was represented by two scales:

Group Teamwork: The extent to which recruits cooperated with one another and worked as a team to get necessary tasks done.

Example Items: (a) In our platoon people cooperated to get things done.

(b) Recruits stressed teamwork and team goals.

Group Support: The extent to which recruits in the platoon tried to make one another feel better when things were going bad and/or provided actual assistance on tasks that did not necessarily require teamwork.

Example Items: (a) Recruits in the platoon trust one another.

(b) Recruits in the platoon lent each other a hand when things got rough.

Self-Improvement Scales

Two multi-item Likert scales developed for Study 2 (2) measured self-improvement:

Self-Esteem: The extent to which the recruit reports that basic training has improved him physically and mentally.

Example Items: (a) I have more self-discipline than when I started training.

(b) After going through boot camp, I believe I can do anything I set my mind to.

Social Skills: The extent to which the recruit feels that basic training has increased his ability to get along with and work with other types of people.

Example Items: (a) As a result of training I've learned to get along with other people much better.

(b) In boot camp I've learned the importance of working together to get things done.

Performance Measures

Performance measures were scores from standard tests routinely administered during BT. These included one academic test administered approximately 2 1/2 weeks into BT and another academic test administered about 7 weeks later at the end of BT. The first test, Practical Examination 1, produced a single score with a maximum value of 100. For the second test, Practical Examination 2, separate scores were recorded for the oral portion and the written portion. The maximum

score for each portion of the second examination was 50. In addition, a total score for the second examination was computed by summing the two tests for comparison to the first examination.

Other performance measures were scores for rifle marksmanship qualification (M16 Score, maximum possible score = 250), and Drill Instructor ratings of general acceptance of Marine Corps standards of behavior and philosophy (Conduct) and of general performance on both tests and routine day-to-day performance during BT (Senior Drill Instructor Subjective Evaluations; hereafter, Subjective Evaluations). The ratings were made at the end of BT; the maximum possible score for each rating was 5.0. Detailed descriptions of each performance measure can be found in the manual of standard operating procedures for training male Marine Corps recruits current when these studies were executed (1).

Health Indicators

Health was measured by number of illness incidents during BT. Illness incidents were identified from health care records kept at the Recruit Training Depot Dispensary. An illness incident was defined as a visit or series of visits for a particular set of presenting complaints with no more than 7 days separating two consecutive visits. If more than 7 days separated two consecutive visits, the visits were scored as separate illness incidents unless there was a specific note that the problem was a continuation of a previous incident.

The diagnosis was recorded for each illness incident. After reviewing the frequency of different types of incidents, this study employed 5 health measures: (a) Total number of incidents, (b) Number of upper respiratory/viral incidents, (c) Number of trauma/injury incidents, (d) Number of "other" health problems, and (e) Number of sprain/strain incidents. Sprains/strains were part of the trauma/injury category, but were singled out for special attention because this type of problem seemed particularly likely to be affected by PT demands.

FMF Success

Success in the Fleet Marine Force (FMF) was scored as follows: 1 = Discharged for behavioral reasons; 2 = Below average rank (E-1 or E-2); 3 = Average rank (E-3); 4 = Above average rank (E-4 or E-5). Men discharged upon completion of active duty requirements (e.g., reservists), for medical reasons, or for miscellaneous nonbehavioral problems were excluded from the FMF success

analyses because their follow-up status was an ambiguous indicator of FMF performance. Data were obtained from computerized records at Headquarters, Marine Corps, 3 to 3-1/2 years after entering BT.

Analysis Procedures

The t-test was used to compare the low and high improvement groups on attitude, leadership, group cohesion, self-improvement, and performance measures. Chi-square tests of association assessed relationships between fitness improvement and health, attrition from BT, and FMF success. Where appropriate, the binomial test was used to test the hypothesis that specific events (e.g., behavioral attrition) were randomly distributed between the low and high improvement groups for specific outcomes (e.g., the number of behavioral attrites in each group) (16, pp. 580-586).

A result was significant if: (a) Group differences were in the same direction in each study. (b) The difference was significant at the $p < .05$ level in at least one study and at the $p < .25$ level in each other study. (c) The pooled significance across studies was $p <$ (.05 / number of variables used to test the hypothesis). These criteria ensured consistent trends across studies, allowed for the fact that even significant associations will produce some nonsignificant findings by chance (17), and kept experiment-wide error probability to 5% or less for each hypothesis (18). The methods of adding logarithms and adding probabilities were used to estimate pooled significance (19).

RESULTS

Hypothesis 1: High fitness improvement will be associated with greater Esprit de Corps.

Hypothesis 1 was supported for attitudes (see Table 2). Each of the 8 attitude comparisons showed more positive attitudes in the high improvement group. The differences met the pooled significance criterion for acceptance of authority and satisfaction with the Marine Corps (pooled significance $p < .001$ for both).

The speculative extension of Hypothesis 1 to include leadership and group cohesion (p. 2) was supported only by the Referent Power results (see Table 2). The high improvement group consistently saw their Drill Instructors as better examples of what a Marine should be (pooled significance, $p < .001$).

TABLE 2

COMPARISON OF ESPRIT DE CORPS INDICATORS
FOR THE LOW AND HIGH FITNESS IMPROVEMENT GROUPS

	STUDY 1				STUDY 2			
	LOW	HIGH	t	SIG.[a]	LOW	HIGH	t	SIG.[a]
ATTITUDES								
Affiliation	3.61	3.96	1.94	.027	4.19	4.22	0.19	.423
Authority	4.67	5.07	2.95	.002	4.83	5.04	2.02	.022
Commitment	5.65	5.72	0.89	.187	5.67	5.78	1.56	.061
Satisfaction	4.99	5.33	2.62	.005	4.69	4.93	1.79	.037
LEADERSHIP								
Leader Structure	4.27	4.74	3.57	.001	4.57	4.25	−2.37	(.010)[b]
Leader Support	4.11	4.65	4.03	.001	3.91	3.87	−0.21	(.419)[b]
Expert Power	5.44	5.71	3.00	.002	5.52	5.56	0.53	.298
Referent Power	3.79	4.44	4.11	.000	4.55	4.91	2.77	.003
GROUP COHESION								
Group Teamwork	3.65	4.02	2.35	.010	3.73	3.49	−1.46	(.071)[b]
Group Support	2.68	3.41	5.49	.001	2.98	2.93	−0.39	(.350)[b]

[a] The indicated significance is one-tailed because there was a directional hypothesis for these analyses.

[b] Group means in this analysis were contrary to the prediction that the high fitness improvement group would have a higher mean.

Hypothesis 2: High fitness improvement will be associated with greater perceived self-improvement in recruits.

Hypothesis 2 was tested only in the second study. The hypothesis was supported for Self-Esteem, but not Social Skills (see Table 3). Informal replication of the Self-Esteem finding was provided by our earlier interviews with recruits, but it must be kept in mind that this finding has not been formally replicated.

TABLE 3

COMPARISON OF SELF-IMPROVEMENT INDICATORS
FOR THE LOW AND HIGH FITNESS IMPROVEMENT GROUPS

	IMPROVEMENT GROUP:			
	LOW	HIGH	t	SIG.[a]
Self-Esteem	5.41	5.66	2.76	.003
Social Skills	5.07	5.06	–0.11	(.455)[b]

NOTE: The self-improvement indicators were only included in Study 2.

[a]The indicated significance is one-tailed because there was a directional hypothesis for these analyses.
[b]Group means in this analysis were contrary to the prediction that the high fitness improvement group would have a higher mean.

Hypothesis 3: Fitness improvement level will be related to BT performance.

The performance hypothesis did not specify the direction of the association because both positive and negative effects were plausible. A general trend toward better performance in the high improvement group was indicated by a higher score in 14 of 15 comparisons (x^2 = 11.27, 1 df, $p < .001$). Despite this consistency, only the oral and total test scores for the second academic examinations were significantly different using the criteria established for the study (see Table 4, p. 12). The pooled significance was $p < .001$ for both.

Hypothesis 4: High fitness improvement will be associated with poorer health in BT.

Data from the second and third studies failed to support Hypothesis 4 even when analysis was restricted to sprain/strain diagnoses. However, a trend toward higher medical attrition in the high improvement group (see below) provided some support for this hypothesis.

TABLE 4

COMPARISON OF HIGH AND LOW IMPROVEMENT GROUPS
ON TRAINING PERFORMANCE MEASURES

	STUDY 1				STUDY 2				STUDY 3			
	LOW	HIGH	t	SIG.	LOW	HIGH	t	SIG.	LOW	HIGH	t	SIG.
Practical Exam 1	88.0	89.0	0.93	.351	87.1	90.8	8.92	.001	87.9	88.1	0.54	.592
M16 Score	202.4	202.2	−0.11	.914	201.8	203.7	2.53	.011	203.8	204.3	0.58	.563
Practical Exam 2:												
Oral	47.4	48.8	3.94	.001	48.3	48.7	3.46	.001	45.1	46.3	5.56	.001
Written	45.4	45.5	0.20	.843	47.2	47.8	3.34	.001	44.8	46.5	9.23	.001
Total	92.7	94.3	2.17	.031	95.5	96.5	4.89	.001	89.9	92.8	9.75	.001

NOTE: Significance levels are 2-tailed because there was no directional prediction for the effects.

Hypothesis 5: Attrition during basic training will be higher in the high improvement group.

Hypothesis 5 was supported using the binomial probability model to test the hypothesis that the distribution of attrites between the low and high improvement groups was equal to that expected by chance. The average overall attrition rate in the low improvement groups was 10.0%; the average for the high improvement groups was 14.5% (see Table 5, pg. 13). Thus, attrition in the high fitness improvement group was 45% higher than in the low improvement platoons. The trend toward excess of attrition in the high improvement group was highly significant in Study 3 ($p < .001$, one-tailed), but only marginally significant in Study 2 ($p < .072$, one-tailed). The pooled significance was $p < .003$, so the overall results clearly supported the hypothesis.

Closer examination of the attrition trends showed that the tendency for the high improvement group to have excess medical attrition was more stable than the tendency to have excess behavioral attrition. Applying the binomial test to the medical attrition data showed a significant difference in Study 2 ($p < .044$) with a marginally significant difference in Study 3 ($p < .112$) (pooled significance, $p < .013$). Applying the same test to the behavioral attrition data produced $p < .283$ for Study 2 and $p < .001$ for Study 3 (pooled significance, $p < .041$). Employing the significance criteria established for this report, the behavioral attrition trend was nonsignificant because the Study 2 probability was $p > .25$.

TABLE 5

PLATOON FITNESS IMPROVEMENT LEVEL AND BASIC TRAINING ATTRITION

STUDY	IMPROVEMENT LEVEL	ATTRITION STATUS		
		BEHAVIORAL	MEDICAL	GRADUATE
2	Low	5.5%	4.7%	89.7%
	High	12.4%	4.6%	83.1%
3	Low	6.0%	3.7%	90.3%
	High	6.7%	5.4%	88.0%

NOTE: Percentages do not sum to 100 due to rounding error.

Hypothesis 6: Graduates from the high improvement group will have greater FMF success than those from the low improvement group.

These analyses were limited to Studies 2 and 3 because of the small sample in Study 1. There were no significant differences in likelihood of discharge or rate of advancement in the FMF.

DISCUSSION AND CONCLUSIONS

Evidently, PT can have side effects during BT as indicated by findings supporting 5 of the 6 hypotheses tested. However, three major issues must be considered before this observation can be translated into guidelines for the design of PT programs. These issues are the mixed positive and negative character of the side effects, the direction of causality for the side effects, and the size of the side effects.

The mixed character of the side effects is evident in the fact that the high improvement groups combined better attitudes toward the Marine Corps and leader-

ship, greater perceived self-improvement, and better academic performance with higher attrition. If the added attrition in the high improvement group represents recruits who could have performed satisfactorily in the FMF, the increased attrition is a negative side effect. However, if the added attrition reflects early identification of recruits who cannot adapt to the the Marine Corps' challenge, the increased attrition would be a positive side effect (1, p. 2-1).

Although the ambiguity of the attrition findings cannot be resolved at this time, several points connected with the attrition results are important for future consideration. First, if the additional attrition represents poor performers or individuals with bad attitudes, the added attrition would improve performance and attitudes. If so, the attrition differences could explain the other differences between the low and high fitness improvement groups. Under these circumstances, the added attrition would be the key side effect for PT because it would explain the other side effects. Second, even if attrition is positive from the organizational perspective, it will be a negative event for most of the individuals discharged. Finally, if high attrition is a negative side effect of high improvement PT programs, low attrition must be regarded as a positive side effect of low improvement PT programs.

The second issue for future consideration is the direction of causality for the associations described in this report. These associations have been interpreted here as indicating that PT caused nonfitness BT outcomes. This focus has been chosen to emphasize the possibility of changing nonfitness BT outcomes by altering PT programs. Other plausible interpretations are possible. For example, positive attitudes at the beginning of BT might contribute to greater effort during training which then produces larger improvements in physical fitness. If PT is not a cause of the nonfitness BT outcomes, changing PT programs will not enhance the effects of BT. Therefore, alternative explanations should be developed and tested now that it has been established that significant associations exist which require explanation.

Longitudinal field studies, particularly field experiments, would be required to test alternative causal hypotheses. Such studies would make it possible to determine whether pre-existing recruit characteristics (e.g., positive attitudes) determine fitness improvement, whether fitness improvement determines how attitudes change, or whether both assertions are true. The results of such studies would be of value no matter what the outcome, because even the alternative explanations would help understand the dynamics producing BT outcomes.

The third issue for further study is whether the side effects of PT are large enough to have practical importance. The differences between the low improvement and high improvement groups were modest in the studies reported here, but methodological factors may have produced this trend. One such factor is the use of insensitive measures for two key dependent variables. Many recruits scored at or near the maximum for several of the attitude and academic performance measures. This indicates that the measures were insensitive to true differences in attitudes and knowledge that occurred in ranges above the upper limit of the scale (20). Because the data trends implied that the high improvement group was higher on these measures, this restriction would affect this group more than the low improvement group. If so, the observed differences underestimated true differences because the high improvement group could not demonstrate its superiority. The development of outcome measures sensitive to the higher ranges of attitudes and knowledge would help resolve this issue.

A second methodological factor that may have affected the size of the associations was the use of an overall fitness measure to classify groups. The problems associated with the use of an overall improvement measure have important implications in connection with the possibility of designing PT programs to promote the full range of BT objectives. Therefore, this point will be developed in some detail.

The use of an overall fitness improvement measure to classify platoons means that platoons with similar overall improvement can differ in important ways. For example, one high improvement platoon may have emphasized strength training while another high improvement platoon emphasized cardiovascular endurance training. Other differences could arise if different methods were used to achieve identical fitness gains. For example, one platoon could have employed infrequent, but high intensity, exercise while another platoon exercised more frequently at a lower intensity.

Platoon differences such as those just described are important if PT side effects depend on the particular type of physical improvement achieved and/or the specific methods employed to produce physical improvement. Extending the second example given above, positive attitude changes may depend on mastering high intensity exercise demands. If so, attitude effects such as those noted in this study would be limited to platoons which emphasized exercise intensity rather than exercising frequently. Analyses comparing groups defined on the basis of overall

improvement will water down the side effects of intense PT by combining high intensity platoons with the low intensity platoons which achieved similar fitness gains but did not produce the attitudinal side effects.

The above example could be extended to many combinations of type, intensity, and duration of exercise, type and amount of fitness improvement, and type of nonfitness BT outcome. The range of possible combinations leads to the conclusion that more detailed assessments of actual patterns of exercise and specific aspects of fitness improvement are needed to accurately estimate the potential for using PT to promote nonfitness BT outcomes. The resulting knowledge would provide a basis for accentuating positive side effects while eliminating negative side effects if the two types of effect depend on different PT elements.

To summarize briefly, the results reported here supported the possibility that PT can affect nonfitness BT outcomes. Having verified this point, more detailed studies of the relationships between specific PT elements and particular nonfitness BT outcomes are needed to determine whether elements of PT programs really cause differences in nonfitness BT outcomes, to determine whether the side effects are large enough to have practical importance, and, if so, to isolate the critical elements of PT for producing positive side effects while avoiding negative side effects. The additional research could provide guidelines for designing PT programs to maximize overall BT outcomes.

REFERENCES

1. Houghton, K.J., Commanding General, Marine Corps Recruit Depot, San Diego, CA. Depot Order P1510.30E of 1 January 1977. Subject: Standing Operating Procedure for Male Recruit Training.

2. Vickers, R.R., Jr., Wallick, M.T. & Hervig, L.K. The Marine Corps basic training experience: Stresses, leadership, and group cohesion as predictors of attitudes, health, and performance. San Diego, CA: Naval Health Research Center, Technical Report 82-28, 1982.

3. Aronson, E. & Mills, J. The effect of severity of initiation on liking for a group. Journal of Abnormal and Social Psychology, 1959, 59, 177-181.

4. Gerard, H.B. & Mathewson, G.C. The effects of severity of initiation on liking for a group: A replication. Journal of Experimental Social Psychology, 1966, 2, 278-287.

5. Houston, B.K., Bloom, L.J., Burish, T.G. & Cummings, E.M. Positive evaluation of stressful experiences. Journal of Personality, 1978, 46, 205-214.

6. Stewart, G.T., Voors, A.W., Jenkins, C.D., Gutekunst, R.R., & Moldow, C.F. Determinants of sickness in Marine recruits. American Journal of Epidemiology, 1969, 89, 254-263.

7. Vickers, R.R. Jr., Haight, M.A., Wallick, M.T. & Conway, T.L. Determinants of sickness in Marine recruits: A replication. San Diego, Ca: Naval Health Research Center, Technical Report 81-15, 1981.

8. Anderson, S., Auquier, A., Hauck, W.W., Oakes, D., Vandaele, W. & Weisberg, H.I. Statistical Methods for Comparative Studies. NY: Wiley, 1980, 235-260.

9. Rogosa, D., Brandt, D. & Simowski, M. A growth curve approach to the measurement of change. Psychological Bulletin, 1982, 92, 726-748.

10. Emerson, J.D. & Strenio, J. Boxplots and batch comparison. In D.C. Hoaglin, F. Mosteller, and J.W. Tukey (eds.), Understanding Robust and Exploratory Data Analysis. NY: Wiley, 1983.

11. Vickers, R.R., Jr. & Ryman, D.H. Development of a questionnaire to measure stress and related concepts in the context of the Marine Corps basic training setting. San Diego, CA: Naval Health Research Center, Technical Report 80-12, 1980.

12. Wallick, M.T., Vickers, R.R., Jr. & Ryman, D.H. Revision of a questionnaire to measure stress and related aspects of basic training. San Diego, CA; Naval Health Research Center, Technical Report 82-19, 1982.

13. Booth, R.F. & Hoiberg, A. Structure and measurement of Marine recruit attitudes. Journal of Applied Psychology, 1974, 59, 236-238.

14. Drucker, E.H. The effects of the merit reward system on morale in basic combat training. Alexandria, VA: Human Resources Research Organization, 1970.

15. Quinn, R.P. & Shepard, L.J. 1972-1973 Quality of Employment Survey: Descriptive Statistics with Comparison Data from the 1969-1970 Survey of Working Conditions. Ann Arbor, MI: Institute for Social Research, University of Michigan, 1974.

16. Hays, W.L. Statistics for Psychologists. NY: Holt, Rinehart, Winston, 1963.

17. Tversky, A. & Kahneman, D. Belief in the law of small numbers. Psychological Bulletin, 1971, 76, 105-110.

18. Dunn, O.J. Multiple comparisons among means. Journal of the American Statistical Association, 1961, 56, 52-64.

19. Rosenthal, R. Combining results of independent studies. Psychological Bulletin, 1978, 85, 185-193.

20. Lord, F.M. Applications of Item Response Theory to Practical Testing Problems. Hillsdale, NJ: Lawrence Erlbaum, 1980.

21. Tatsuoka, M.M. Multivariate Analysis. NY: Wiley, 1971.

22. Nunnally, J.C. Psychometric Theory (Second Edition). NY: Wiley, 1978.

23. Nie, N.H., Hull, C.H., Jenkins, J.G., Steinbrenner, K. & Bent, D.H. Statistical Package for the Social Sciences. NY: McGraw-Hill, 1975.

24. Cohen, J. Statistical Power Analysis for the Behavioral Sciences. NY: Academic Press, 1969.

APPENDIX A

DEFINING FITNESS IMPROVEMENT GROUPS

Testing the research hypotheses required that platoons be characterized in terms of fitness improvement during BT. It was therefore necessary to estimate changes in fitness from early in BT to the end of BT. Because change measurement is a complex statistical problem, a series of analyses were completed to ensure that fitness improvement was measured properly and that the resulting platoon classifications provided legitimate tests of the hypotheses described in the introduction.

Sample Comparisons. Initial analyses compared the PFT score distributions for the three samples to determine whether the samples could be regarded as deriving from a single population. If not, the three studies could not reasonably be treated as replications of one another. Results showed:

(a) The average PFT scores varied substantially for both tests, but variability around the means was stable across studies (see Table A-1).

(b) The average difference between PFT2 and PFT1 ranged from 41.28 in Study 3 to 48.02 in Study 2. The variation was statistically significant ($F = 10.00$, $p < .001$), but accounted for only 0.5% of the overall variation in PFT2-PFT1 differences. Also, the 7-point range for differences was substantially smaller than the 21 point range for PFT1 and the 17-point range for PFT2. Thus, amount of change was roughly comparable over studies.

(c) The regression lines predicting PFT2 from PFT1 were highly similar across studies (see Table A-1). Each regression coefficient fell within the 50% confidence intervals for the other two studies.

TABLE A-1

PHYSICAL FITNESS TEST DATA

COHORT	PFT1		PFT2			REGRESSION COEFFICIENTS	
	MEAN	S. D.	MEAN	S. D.	r_{12}	b	a
Study 1	173.50	41.35	219.99	38.17	.618	.570	121.46
Study 2	187.16	43.75	235.14	38.18	.687	.599	123.00
Study 3	194.79	.43.91	236.07	35.45	.720	.581	122.80

NOTE: The regression coefficients produce predictions of PFT2 based on PFT1 as follows: Predicted PFT2 = a + b(PFT1).

The results for this initial analysis of the fitness scores indicated sufficient similarity to treat the samples as replications. This conclusion was appropriate given the emphasis on fitness improvement. The sample variations in PFT2-PFT1 differences and the regression lines required for computing residualized gains did not appear sufficiently large to regard the samples as noncomparable.

Platoon Differences in Fitness Improvement. The next set of analyses confirmed that fitness improvement varied significantly across BT platoons. This confirmation was a logical requirement for asserting that platoons truly had different PT programs. Comparing programs which differed by only chance amounts would not provide a meaningful test of our hypotheses. Analyses showed:

(a) Fitness improvements varied significantly across platoons. This was true of both the simple difference between PFT2 and PFT1 and the residualized gain score described in the Method section (see Table A-2). Note that the platoon differences in the two improvement measures were much more substantial than the platoon differences in either initial fitness level or final fitness level, as indicated by the larger eta coefficient. Note also that the significant differences for the gain scores reduces the likelihood that the results derived from statistical artifacts arising from the presence of PFT1 differences (e.g., regression to the mean).

TABLE A-2

RESULTS OF ONE-WAY ANALYSES OF VARIANCE
COMPARING INDIVIDUAL PLATOONS
ON INITIAL FITNESS, FINAL FITNESS, AND FITNESS IMPROVEMENT MEASURES

	PFT 1			PFT 2			GAIN			DIFFERENCE		
	eta	F	SIG.	eta	F	SIG.	eta	F	SIG.	eta	F	SIG.
Study 1	.33	1.39	.106	.43	2.69	.001	.51	4.18	.001	.48	3.49	.001
Study 2	.28	3.52	.000	.36	6.49	.000	.47	12.08	.000	.45	10.61	.001
Study 3	.33	2.75	.000	.43	2.05	.000	.51	6.12	.000	.36	7.29	.001

NOTE: See Method section for description of the number of platoons and recruits in each study.

(b) The fitness improvement measures were highly correlated across platoons (see Table A-3). Thus, classifications could be developed which minimized important problems in the assessment of change by combining the four potential classification criteria (see pp. 3-4).

TABLE A-3

RANK ORDER CORRELATIONS BETWEEN DIFFERENT POSSIBLE CRITERIA FOR ESTABLISHING PHYSICAL TRAINING INTENSITY LEVELS

Study 1 (n = 26)				
Mean for Differences	--			
Median for Differences	.91	--		
Mean for Gain	.95	.90	--	
Median for Gain	.90	.93	.95	--
Study 2 (n = 40)				
Mean for Differences	--			
Median for Differences	.97	--		
Mean for Gain	.91	.88	--	
Median for Gain	.92	.89	.98	--
Study 3 (n = 37)				
Mean for Differences	--			
Median for Differences	.97	--		
Mean for Gain	.93	.91	--	
Median for Gain	.89	.89	.96	--

Tests of Regression Assumptions for Gain Scores. Having established that individual platoons differed sufficiently to provide a reasonable basis for defining high and low improvement groups, the platoons in each of the three samples were divided into (a) those falling in the upper 40% of both the difference and gain averages, (b) those falling in the lower 40% on both, and (c) all other platoons. Further analyses then considered several possibilities which might have affected the validity of the gain scores as bases for classification:

(a) Gain scores accurately estimated group differences only if the regression lines relating PFT2 to PFT1 were comparable within groups (21, pp. 40-50). Analysis of covariance showed non-parallel regression lines for the first two studies, but not the third (see Table A-4). The statistically significant differences were small. Over the three studies, an average of 45% of the variation in PFT2 scores was predictable from the PFT1 scores based on the overall sample regression line. Replacing the overall sample regression line with 3 within group regres-

sion lines in each study would have predicted an average of 0.5% more of the PFT2 variance. Thus, the similarity across improvement levels was much more substantial than the differences.

TABLE A-4

COMPARISON OF REGRESSION LINES
FOR DIFFERENT FITNESS IMPROVEMENT LEVELS

	TEST FOR NONPARALLELISM: F	SIG.	eta^2 *	WITHIN GROUP REGRESSION SLOPES: LOW	MEDIUM	HIGH
Study 1	6.35	.002	.019	.650	.626	.527
Study 2	4.53	.011	.008	.666	.637	.557
Study 3	1.57	.208	.001	.586	.625	.596

*eta^2 is the proportion of total variance explained by the differences between the within group regression lines.

(b) The small within group differences in regression slopes that were observed probably arose from having an upper limit for fitness test scores. Examination of the scatterplots for the high intensity groups indicated that more recruits in this group were at or near the 300-point maximum for the test. This fact could mean that the PFT2 scores for this group were affected more than the other groups by a restriction of range which would effectively reduce the magnitude of the regression slope (22, pp. 140-141).

(c) When measuring change, controls must be introduced to adjust for the effects of variables which might bias change estimates (8,9). Bias would arise if there were variables which were (i) correlated with fitness improvement and (ii) nonrandomly distributed across the fitness improvement g ps. Exploratory analyses showed that no sociodemographic va ables were significant sources of bias in the fitness improvement estimates. Although high improvement platoons consistently had a below average proportion of whites, the correlation between race and fitness improvement was minimal. For the gain scores, the weighted average correlation adjusting for study differences in sample size was r = .03 (range: r = -.11 to r = .06). The comparable figures for difference scores were r = -.01 (range: r = -.11 to r = .00). No other sociodemographic variable was nonrandomly distributed across the fitness improvement groups in each of the studies.

Effectiveness of Group Classification. A final set of analyses examined the clarity and consistency of discrimination between high and low intensity platoons. Especially in an initial study such as the present, it is desirable that the groups being compared differ widely on the classification variable(s). Without substantial differences, comparisons between the so-called "high improvement" and "low improvement" groups would be meaningless. Although the classification procedures employed for defining the high improvement and low improvement groups guaranteed at least some difference, it was of interest to determine how substantial the differences were. Results were:

(a) In all three samples, the high and low improvement groups differed substantially ($p < .001$) on both the difference scores and the gain scores. The gain score differences confirmed that the high improvement group exceeded the low improvement group taking any initial group fitness differences into account.

(b) Pairwise comparisons between individual high intensity platoons and individual low intensity platoons were made using the least significant difference test for multiple comparisons with a modification to adjust for differences in group sizes (23). The significance criterion for pairwise comparisons was based on a 10% significance level for the overall set of comparisons. The proportion of pairwise comparisons between high and low intensity platoons exceeding the criterion value were: (i) Study 1, 84.7% for gain scores; 18.1% for difference scores. (ii) Study 2, 96.5% for gain comparisons; 97.2% for difference scores. (iii) Study 3, 95.1% for gain scores; 94.4% for difference scores.

These findings showed good discrimination between the high and low improvement groups except in Study 1. However, the typical difference between platoons in Study 1 was roughly as large as that in the other two studies. Therefore, the smaller size of the samples representing each platoon was the reason for the lower frequency of significant between platoon differences (24). Overall, the discrimination between high improvement and low improvement platoons was adequate to test our hypotheses.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER 83-37	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) SIDE EFFECTS OF PHYSICAL TRAINING: ASSOCIATION OF FITNESS IMPROVEMENT TO ESPRIT DE CORPS, PERFORMANCE, HEALTH, AND ATTRITION IN MARINE CORPS BASIC TRAINING		5. TYPE OF REPORT & PERIOD COVERED Interim 6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Ross R. Vickers, Jr.		8. CONTRACT OR GRANT NUMBER(s)
9. PERFORMING ORGANIZATION NAME AND ADDRESS Naval Health Research Center P.O. Box 85122 San Diego, California 92138-9174		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS M0096-PN.001-1035
11. CONTROLLING OFFICE NAME AND ADDRESS Naval Medical Research and Development Command Naval Medical Command, National Capital Region Bethesda, Maryland 20814		12. REPORT DATE December 1983 13. NUMBER OF PAGES 29
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office) Commander, Naval Medical Command Department of the Navy Washington, D.C. 20372		15. SECURITY CLASS. (of this report) UNCLASSIFIED 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

Approved for public release; distribution unlimited.

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Physical training	Health
Military recruits	Attrition
Attitudes	Performance
Leadership	Self-esteem
Group Cohesion	

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Physical training (PT) is a significant element of Marine Corps basic training (BT) which may affect nonfitness BT outcomes in addition to improving fitness. If so, PT side effects should be considered when designing and evaluating PT programs.

This study capitalized on naturally occurring platoon differences in fitness outcomes to classify BT platoons into relatively low and high improvement groups. The high improvement group had better attitudes toward the Marine

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

S/N 0102-LF-014-6601

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

20. Abstract (continued)

Corps, greater perceived self-improvement, and performed better during BT. The groups had comparable illness rates during BT, but there was more medical attrition and overall attrition in the high improvement group. The groups did not differ with regard to Fleet Marine Force promotion or attrition rates.

These findings demonstrate that PT programs may indeed influence nonfitness BT outcomes. Additional research is needed to verify that PT differences actually cause nonfitness BT outcomes, to determine whether these side effects are large enough to be of practical importance, and to identify specific aspects of PT programs that are important for promoting nonfitness BT objectives.